The Analytical Information Markup Language Project

</ 

Gary W. Kramer

Biochemical Science Division
National Institute of Standards and Technology

NIST

Folks, we got TROUBLE, right here in Science City. It starts with T, and it rhymes with D, and it stands for DATA…

- Can't move it
  - From instrument to instrument
  - From instrument to application
  - From application to application
- Can't interconvert it
- Can't find all its parts
- Can't look at it easily
- Can't use it easily with modern computing and networking technologies

NIST

There is crisis in archiving and retrieving data

- We're drowning in data, yet increasingly we cannot find our stuff because of
  - Incredible Success of Hyphenated and Multiplexed Analytical Techniques
  - Moore's Law
- Data mining can increase the value of archived data, but…
- Regulatory agencies are now demanding extraordinarily long data retention
- In terms of retrieving archived data, we were better off with paper forty years ago before the arrival of lab computers…
- We can still read the data in Newton's notebooks today; will folks be able to read ours in 100, 20, or even 5 years?

NIST

What Are the AnIML Standards?

- The AnIML Standards are projects of ASTM Committee E13 on Spectroscopy and Separation Science, Subcommittee E13.15 on Analytical Data

- AnIML is a markup language based on Extensible Markup Language (XML)

NIST

The Analytical Information Markup Language (AnIML)

- has been developed expressly for analytical chemistry result data and metadata
- defines a data format for representing all types of result data and metadata
- provides a system for identifying content (data elements and attributes) with labels (tags)
- provides a model and the means for defining the structure, content, and semantics of result data and metadata in documents
- expresses shared vocabularies
- allows machines to carry out rules made by people

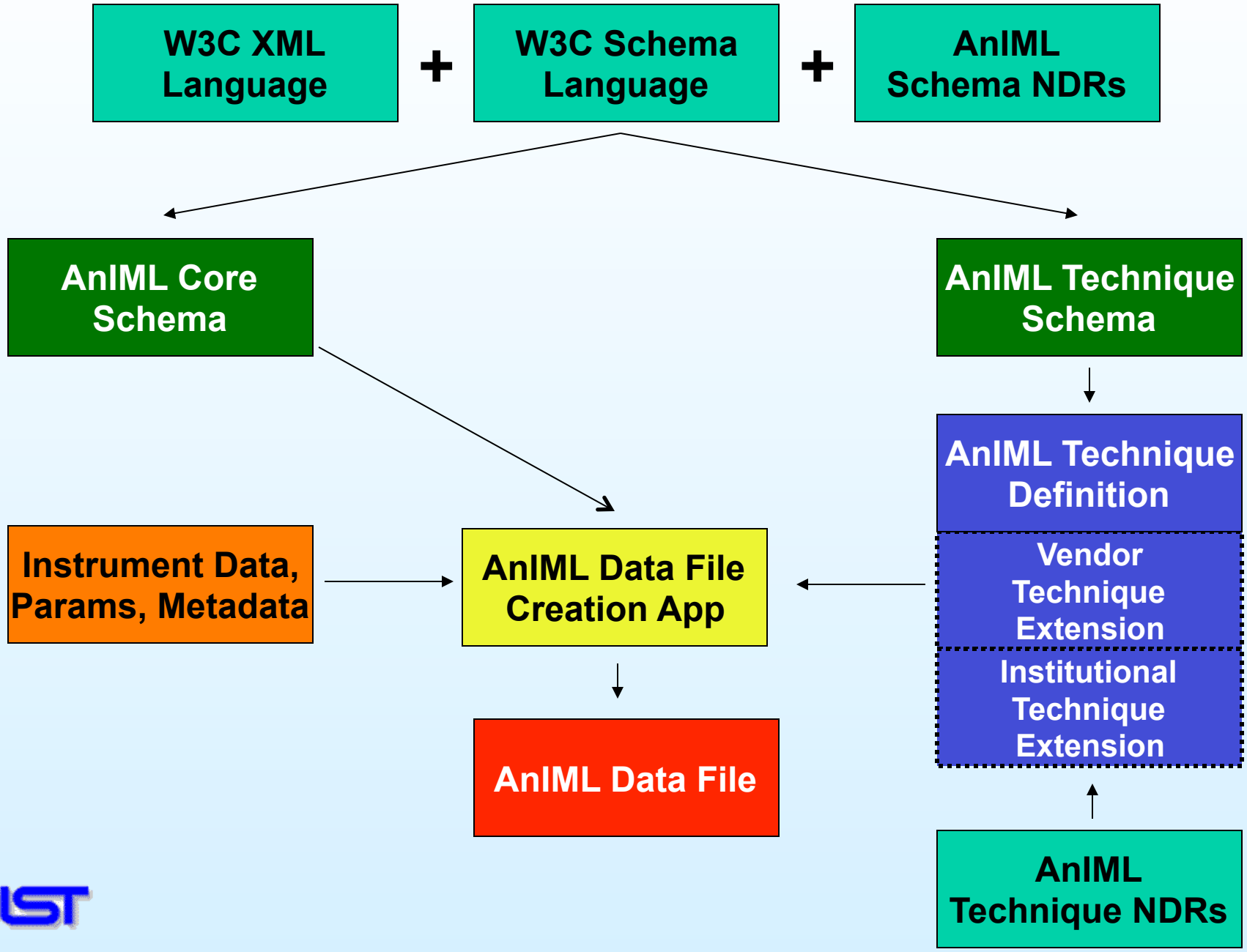
NIST

AnIML Components

- AnIML Core Schema
  - One Schema
  - Maintained by ASTM E13.15
- AnIML Technique Schema
  - One Schema
  - Maintained by ASTM E13.15
- AnIML Technique Definition Documents
  - One or More Instance Documents per Technique
  - Maintained by ASTM E13 or Appropriate Domain Expert Organization
- AnIML Technique Definition Extensions
  - One Instance Document per Technique Created from the Appropriate Base Technique Definition
  - Maintained by Vendor, Organization, User, or Whomever Extends the Technique
- AnIML Result Data Files

NIST

Basis for Creating an AnIML Data File



Some AnIML Files are Simple

Some AnIML Files are Complex

Key Features of AnIML

- Extensible
- Can Accommodate ANY Type of Data
- Verifiable Using Digital Signatures
- Traceable through Audit Trails
- Validateable Against Data Format Rules, Data Structure Rules, & other Data Requisites

NIST

AnIML: Where We're At for Version 1.0

- AnIML Core Schema
  - Complete
  - Frozen
  - Internally Documented
- AnIML Schema NDRs
  - Complete
- AnIML Technique Schema
  - Complete
  - Frozen
  - Internally Documented
- AnIML Technique NDRs
  - In Progress
- AnIML Technique Exemplar
  - In Progress

NIST

AnIML: Where We're At for Version 1.0

- AnIML Technique Definitions
  - UV-Vis Technique Definition – Complete
  - Chromatography Technique Definition – Complete
  - Mass Spec Technique Definition – Started
  - 1D NMR Technique Definition – Started
  - IR Technique Definition – Started
  - Point Detector Technique Definitions
    - GC Detectors (5) – Complete
    - LC Detectors – Started
  - AnIML Misc. Technique Definitions
    - Indexing – Complete
    - Cross-Cutting – Not Started
    - Trace Forming – Not Started
    - Peak Table – Not Started

NIST

AnIML: Where We're At for Version 1.0

- AnIML Applications and Tools
  - AnIML Data Viewers – Several Complete
  - AnIML Extensible Validator – Complete
  - Simple AnIML File Writer for Agilent 8453 UV/Vis – Complete
  - JCAMP-DX to AnIML Data Converters – Several Complete
  - XSLT Documentation Extractor for Schemas – Complete
  - Documentation Extractor for Technique Definitions – Complete

NIST

AnIML Documents and Standards

- AnIML Standard Practice Document – Nearly Complete
- AnIML Core and Technique Schema Standard Specification and Guide Document
- AnIML Standard Technique Definition Guide Document
- Individual Technique Standard Specification and Guide Documents
  - AnIML UV-Vis Technique
  - AnIML IR Technique
  - AnIML Chromatography Technique
  - AnIML Mass Spec Technique
  - AnIML 1D NMR
  - AnIML Misc. Techniques
    - Indexing
    - Point Detectors
    - Cross-Cutting
    - Trace Forming
    - Peak Table

NIST

Expediting the Completion of AnIML 1.0

- Current volunteer effort is too slow
- There is still too much to do
- Need paid help to develop components and documentation
- AnIML 1.0 Documentation Expedition Project
  - Funded by
    - Agilent Technologies
    - Amgen
    - Glaxo SmithKline
    - PerkinElmer
    - Waters Corporation
  - Administered by ASTM

NIST

Want to Know More about AnIML?

- AnIMLWebsite
  - http://animl.sourceforge.net
  - http://www.animl.org
- PittCon 2012
  - SEDD & AnIML Networking Session - Monday AM
  - AnIML Workshop – Thursday PM
- SLAS 2012
  - 2:00 PM Today Room 1AB Burkhard Schäfer
  - BSSN White Paper

NIST

Want to Help Develop AnIML?

- Contact
  - Gary Kramer, NIST
  - +1-301-975-4132
  - gary.kramer@nist.gov

  - Burkhard Schäfer, BSSN
  - +1-888-674-0047
  - +49-6136-75-27-60
  - bschaefer@bssn-software.com
- Webex Virtual Meetings
  - Monthly – 10:30 AM EST
  - ASTM-sponsored

NIST

Acknowledgements

Mark Bean – GlaxoSmithKline

Maren Fiege – Waters Informatics GmbH

Burkhard Schäfer – BSSN

Stuart Chalk – University of North Florida

Dale O'Neill & Michael Leiber – Agilent Technologies

Adam Patkin – PerkinElmer

Jamie McQuay – Scimatic Software

Peter Linstrom – NIST

Tony Davies – IUPAC Subcommittee on Electronic Data Standards

Reinhold Schäfer – Fachhochschule Wiesbaden

- Alex Rühl
- Martin Peschke
- Aykut Arslan
- Dominick Pötz
- Anh Dao Nguyen
- Alex Roth
- Ronnie Jopp
- Patrick Gleichmann
- Kordian Placzek
- Frank Masur
- Dennis Backhaus

NIST